



UNIVERSITÀ  
DEGLI STUDI  
FIRENZE

# FLORE

## Repository istituzionale dell'Università degli Studi di Firenze

### On the twist condition and c-monotone transport plans.

Questa è la Versione finale referata (Post print/Accepted manuscript) della seguente pubblicazione:

*Original Citation:*

On the twist condition and c-monotone transport plans / Champion, Thierry; De Pascale, Luigi; - In: DISCRETE AND CONTINUOUS DYNAMICAL SYSTEMS. - ISSN 1078-0947. - STAMPA. - 34:(2014), pp. 1339-1353. [10.3934/dcds.2014.34.1339]

*Availability:*

This version is available at: 2158/1070953 since: 2017-07-03T12:00:05Z

*Published version:*

DOI: 10.3934/dcds.2014.34.1339

*Terms of use:*

Open Access

La pubblicazione è resa disponibile sotto le norme e i termini della licenza di deposito, secondo quanto stabilito dalla Policy per l'accesso aperto dell'Università degli Studi di Firenze (<https://www.sba.unifi.it/upload/policy-oa-2016-1.pdf>)

*Publisher copyright claim:*

(Article begins on next page)

# On the twist condition and $c$ -monotone transport plans

Thierry Champion

*Institut de Mathématiques de Toulon et du Var, U.F.R. des Sciences et Techniques, Université du Sud  
Toulon-Var, Avenue de l'Université, BP 20132, 83957 La Garde cedex, FRANCE*

Luigi De Pascale\*

*Dipartimento di Matematica Applicata, Università di Pisa, Via Buonarroti 1/c, 56127 Pisa, ITALY*

---

## Abstract

We consider optimal transport problems in  $\mathbb{R}^d$  as well as on manifolds for cost functions  $c$  which satisfy a nonsmooth version of the classical Left Twist condition (i.e. the invertibility of the gradient in the first variable). Under the classical assumption that the initial measure does not give mass to sets with  $\sigma$ -finite  $\mathcal{H}^{d-1}$  measure, we provide a short and self-contained proof of the fact that any optimal transport plan as well as any feasible transport plan satisfying a  $c$ -monotonicity assumption is induced by a transport map. We also show that the usual costs induced by Tonelli Lagrangians satisfy the Nonsmooth Left Twist condition we propose.

*Keywords:* Monge-Kantorovich problem, optimal transport problem, cyclical monotonicity, Tonelli Lagrangian.

*2000 MSC:* 49Q20, 49K30, 49J45

---

## 1. Introduction

Consider the optimal transport problem

$$(P) \quad \inf \left\{ \int_{M \times N} c(x, y) d\gamma(x, y) : \gamma \in \Pi(\mu, \nu) \right\}$$

where  $M$  is a second countable  $\mathcal{C}^1$  manifold,  $N$  is a Polish space,  $c : M \times N \rightarrow \mathbb{R}$  is a cost function,  $\mu$  and  $\nu$  are Borel probabilities respectively over  $M$  and  $N$ , and  $\Pi(\mu, \nu)$  denotes the set of transport plans from  $\mu$  to  $\nu$ , i.e. the set of Borel probabilities  $\gamma$  on  $M \times N$  with first marginal  $\pi_x^* \gamma$  equal to  $\mu$  and second marginal  $\pi_y^* \gamma$  equal to  $\nu$ . In this work, we are interested in identifying conditions on the cost function  $c$  and the measure  $\mu$  to ensure the existence of

---

\*Corresponding author

Email addresses: `champion@univ-tln.fr` (Thierry Champion), `depascal@dm.unipi.it` (Luigi De Pascale)

an optimal transport plan  $\gamma$  for  $(\mathbf{P})$  which is induced by a transport map, that is for which there exists a  $\mu$ -measurable function  $T : M \rightarrow N$  such that  $\gamma = (id \times T)_\# \mu$ .

This problem has been extensively studied in the past twenty years, and one may distinguish two main types of sufficient conditions for the existence of an optimal transport plan for  $(\mathbf{P})$  induced by a map, depending on whether or not the cost function  $c$  satisfies the following regularity assumption :

$$(\mathbf{LT}) \quad \text{for all } x \in M, y \mapsto \frac{\partial c}{\partial x}(x, y) \text{ is injective on its domain,}$$

where  $(\mathbf{LT})$  stands for “left twist” (we refer to Fathi and Figalli [1] for comments on this terminology). For example, in the special case where  $M = N = \mathbb{R}^d$  and  $c(x, y) = \|y - x\|^p$ , with  $p > 0$  and  $\|\cdot\|$  denoting the Euclidian norm, the assumption  $(\mathbf{LT})$  holds whenever  $p \neq 1$  but fails for  $p = 1$ . Let us briefly describe the corresponding existence results for those particular instances of problem  $(\mathbf{P})$ . First in the regular case where  $(\mathbf{LT})$  holds, *i.e.* for  $p \neq 1$ , one obtains that the unique solution of  $(\mathbf{P})$  is indeed induced by a map under the general hypothesis that  $\mu$  does not give mass to sets with  $\sigma$ -finite  $\mathcal{H}^{d-1}$  measure, see Brenier [2] and R  schendorf and Rachev [3] for the case  $p = 2$  and Caffarelli [4], Gangbo and McCann [5, 6] and R  schendorf [7] for the general case  $p \neq 1$ . On the other hand the case  $p = 1$ , which corresponds to the relaxation proposed by Kantorovitch [8, 9] for the optimal mass transport problem originally studied by Monge [10], this existence issue reveals quite involved and one only obtains that some particular solution of  $(\mathbf{P})$  is indeed induced by a map under the hypothesis that  $\mu$  is absolutely continuous with respect to the Lebesgue measure  $\mathcal{L}^d$ , see Evans and Gangbo [11] with the further assumption that  $\mu$  and  $\nu$  have Lipschitz densities with respect to  $\mathcal{L}^d$ , and Ambrosio and Pratelli [12], Caffarelli, Feldman and McCann [13], Caravenna [14], Champion and De Pascale [15, 16], and Trudinger and Wang [17] for the general case on  $\mu$ . In that case the problem  $(\mathbf{P})$  may have solutions which are not induced by a transport map even when both measures  $\mu$  and  $\nu$  are absolutely continuous with respect to the Lebesgue measure  $\mathcal{L}^d$ , or may have no solution induced by a transport map in the case where  $\mu$  is not absolutely continuous with respect to the Lebesgue measure  $\mathcal{L}^d$  (see Ambrosio and Pratelli [12]).

In the present work we shall concentrate on the regular case where the assumption  $(\mathbf{LT})$  holds. In this framework and under the hypothesis that the admissible plan  $\gamma$  is concentrated on a  $c$ -cyclically monotone set (which amounts to  $\gamma$  being optimal for  $(\mathbf{P})$  when the value is finite, see §4 below), the fact that the plan  $\gamma$  is induced by a transport map has been proven by many authors under various regularity hypotheses. When  $M = \mathbb{R}^d$ , this was in particular obtained by Caffarelli [4], Gangbo and McCann [5, 6] and R  schendorf [7] in the case  $c(x, y) = h(y - x)$  (see also Villani [18] for a presentation), while the general form for the cost function  $c$  was investigated in Carlier [19], Levin [20]. For the case where  $(M, d)$  is a Riemaniannian manifold, this was studied for the square distance cost  $c = d^2$  by Cordero-Erausquin [21] (in the special case of the flat torus) and McCann [22], and the general case where  $c$  is induced by a Lagrangian was studied by Bernard and Buffoni [23], Fathi and Figalli [1] and Figalli [24, 25]. While all these works heavily use the fact that the plan  $\gamma$  is concentrated on a  $c$ -cyclically monotone set, our paper is innovative in that it applies to any plan  $\gamma$  which is supported on a  $c$ -monotone set (see section 4 below) : to obtain such a result we have to strengthen the hypothesis  $(\mathbf{LT})$  on the cost function  $c$  (see the Nonsmooth Left

Twist condition **(NLT)** in section 2 below) but we prove that the most common cost functions that satisfy **(LT)** also satisfy this stronger condition. Surprisingly enough our result holds true with the same requirement that the initial measure  $\mu$  shall give zero mass to sets with  $\sigma$ -finite  $\mathcal{H}^{d-1}$  measure (in fact we slightly weaken this hypothesis). Our approach follows the method of proof designed by Champion, De Pascale and Juutinen [26] for the study of the  $\infty$ -Wasserstein distance problem and generalized by Champion and De Pascale [16, 27] to deal with the classical Monge Transportation problem. The main features of our proof is that it relies on local arguments and make no use of the regularity theory for  $c$ -convex functions; see the discussion following the proof of Theorem 4.3 below.

The paper is organized as follows : in sections §2 and §3 we present the assumptions we shall make on the cost function  $c$  and the initial measure  $\mu$  respectively. In §3 we also deduce some regularity property for any admissible plan  $\gamma \in \Pi(\mu, \nu)$ . Section 4 is devoted to our main result and some comments, and we prove that this result applies to the most usual costs induced by Lagrangians in section 5.

## 2. Hypotheses on the cost function

We hereafter describe the common regularity assumptions that we shall impose on the cost function  $c$ , and in particular we introduce our refined version of the left twist condition **(LT)** which is the basis of our main result Theorem 4.3. Since our method of proof is local, we first state those assumptions in  $U \times N$ , where  $U$  is an open subset of  $\mathbb{R}^d$  and  $N$  is a Polish space, and we shall then briefly discuss the manifold setting, the specific case of costs induced by a Tonelli Lagrangian being treated in section §5 below. A more detailed presentation of the following notions (except for the nonsmooth left twist condition **(NLT)** below) and related properties in the context of optimal transportation may be found for example in chapter 10 of Villani [28] or the appendix of Fathi and Figalli [1].

**Definition 2.1.** *The function  $c : U \times N \rightarrow \mathbb{R}$  is superdifferentiable in  $x$  at  $(x_0, y_0)$  if there exist  $p_0 \in \mathbb{R}^d$  and a function  $\eta_0 : \mathbb{R}^+ \rightarrow \mathbb{R}$  such that*

$$\forall x \in U, \quad c(x, y_0) \leq c(x_0, y_0) + p_0 \cdot (x - x_0) + \eta_0(\|x - x_0\|)\|x - x_0\| \quad (1)$$

*with  $\eta_0(r) \rightarrow 0$  as  $r \rightarrow 0$ , where  $\cdot$  denotes the canonical scalar product on  $\mathbb{R}^d$ . The convex set of all such vectors  $p_0$  is the superdifferential in  $x$  of  $c$  at  $(x_0, y_0)$  and is denoted by  $\nabla_x^+ c(x_0, y_0)$ . The likewise notions of subdifferentiability and subdifferential  $\nabla_x^- c(x_0, y_0)$  are obtained by reversing the inequality in (1).*

**Remark 2.2.** *Note that a function  $c$  is differentiable with respect to  $x$  at  $(x_0, y_0)$  if and only if it is both sub- and superdifferentiable with  $\nabla^+ c(x_0, y_0) = \nabla^- c(x_0, y_0) = \{\nabla_x c(x_0, y_0)\}$  (we refer to Proposition 10.7 of Villani [28]).*

We shall need the usual following refined notion of local regularity.

**Definition 2.3.** *The function  $c : U \times V \rightarrow \mathbb{R}$  is locally superdifferentiable in  $x$  at  $(x_0, y_0)$  locally in  $y$  if there exists a neighborhood  $W$  of  $(x_0, y_0)$  such that  $c$  is superdifferentiable in*

$x$  at any  $(x, y) \in W$  and there exists a continuous function  $\eta : \mathbb{R}^+ \rightarrow \mathbb{R}$  such that for all  $(x, y) \in W$  :

$$\forall x', \forall p \in \nabla_x^+ c(x, y), \quad c(x', y) \leq c(x, y) + p \cdot (x' - x) + \eta(\|x' - x\|)\|x' - x\|$$

with  $\eta(r) \rightarrow 0$  as  $r \rightarrow 0$ . The corresponding notion of local subdifferentiability is defined likewise.

Notice that in the above local notion of superdifferentiability it is assumed that the function  $\eta$  involved in (1) does not depend on  $(x, y)$  locally around  $(x_0, y_0)$ .

**Remark 2.4.** It is to be noted that when  $c : U \times V \rightarrow \mathbb{R}$  is locally superdifferentiable (resp. subdifferentiable) in  $x$  at  $(x_0, y_0)$  locally in  $y$ , then it is differentiable in  $x$  at  $(x_0, y_0)$  whenever  $\nabla_x^+ c(x_0, y_0)$  is reduced to a singleton (see Theorem 10.8(iii) in Villani [28]).

**Definition 2.5.** A set-valued map  $F : U \times N \rightrightarrows \mathbb{R}^d$  is locally bounded if for any  $(x, y) \in U \times N$  there exists a neighborhood  $W$  of  $(x, y)$  and a compact set  $K$  such that  $F(x, y) \subset K$  for all  $(x, y) \in W$ .

The following result is a direct consequence of the above definitions, and states that under mild assumptions the superdifferential (resp. subdifferential) set-valued map has closed values and closed graph.

**Lemma 2.6.** Assume  $c : U \times N \rightarrow \mathbb{R}$  is continuous, locally superdifferentiable in  $x$  locally in  $y$  on  $U \times N$ , and that  $(x, y) \mapsto \nabla_x^+ c(x, y)$  is locally bounded in  $U \times N$ . Then whenever  $(x_n, y_n)_{n \geq 0}$  converges to  $(x_\infty, y_\infty)$  and  $p_n \in \nabla_x^+ c(x_n, y_n)$  for any  $n$ , it holds that  $(p_n)_{n \geq 0}$  is bounded and any of its cluster points  $p_\infty$  belongs to  $\nabla_x^+ c(x_\infty, y_\infty)$ . A similar statement holds for the subdifferential.

As already mentioned, the proof of our main result is based on a local argument, so that for our purpose all the previous notions may be generalized through local charts to a cost  $c : U \times N \rightarrow \mathbb{R}$  where  $U$  is an open set of a  $\mathcal{C}^1$  manifold  $M$ . Indeed, if  $(V, \phi)$  and  $(W, \psi)$  are two local charts around  $\mathbf{x}_0 \in U$ , then  $c \circ (\phi^{-1} \times id)$  is locally superdifferentiable with respect to the first variable  $x$  locally in  $y$  at  $(\phi(\mathbf{x}_0), y_0)$  if and only if the same holds for  $c \circ (\psi^{-1} \times id)$  at  $(\psi(\mathbf{x}_0), y_0)$ , and in this case one has

$$\nabla_x^+ (c \circ (\phi^{-1} \times id))(\phi(\mathbf{x}), y) = {}^t D(\psi \circ \phi^{-1}) [\nabla_x^+ (c \circ (\psi^{-1} \times id))(\psi(\mathbf{x}), y)] \quad (2)$$

for any  $(\mathbf{x}, y) \in (V \cap W) \times N$  : this directly follows from the fact that  $\psi \circ \phi^{-1}$  is a  $\mathcal{C}^1$  diffeomorphism on  $\phi(V \cap W)$ . Note that equation (2) uniquely characterizes the following subset of the cotangent space  $T_{\mathbf{x}_0}^* M$  :

$$\nabla_{\mathbf{x}}^+ c(\mathbf{x}_0, y_0) = \phi^* (\nabla_x^+ (c \circ (\phi^{-1} \times id))(\phi(\mathbf{x}_0), y_0)). \quad (3)$$

Indeed such a subset does not depend on the local coordinate and is, by definition, the superdifferential of  $c$  with respect to  $\mathbf{x}$  at  $(\mathbf{x}_0, y_0)$ . The more specific case where  $M$  is endowed with a Riemannian structure is treated in section §5 below.

We finally introduce the refined twist condition.

**Definition 2.7.** Assume  $c : U \times N \rightarrow \mathbb{R}$  is superdifferentiable in  $x$  on  $U \times N$ , we say that  $c$  satisfies the Nonsmooth Left Twist condition whenever it holds

(NLT) for all  $x \in U$ , the set-valued map  $y \mapsto \nabla_x^+ c(x, y)$  is injective.

**Remark 2.8.** We use the same terminology of “Nonsmooth Left Twist condition” when  $c$  is subdifferentiable, with  $\nabla_x^- c(x, y)$  replacing  $\nabla_x^+ c(x, y)$  : it indeed follows from Remark 2.2 that these two conditions coincide with (LT) when  $c$  is both sub- and superdifferentiable.

Let us comment briefly on the above definition. First, the conditions (LT) and (NLT) obviously coincide in the regular case when  $c$  is differentiable in  $x$  on its domain. This is for example the case for the usual  $l^p$  costs  $c(x, y) = \|y - x\|^p$ , which do satisfy both conditions on  $\mathbb{R}^d \times \mathbb{R}^d$  for  $p$  positive and  $p \neq 1$ . In fact the main reason for introducing this nonsmooth version of the left twist condition is to address the case of non-regular costs  $c$  which do naturally appear in the case where  $M$  is a manifold. In particular, if one considers the case of a smooth Riemannian manifold  $(M, g)$  with corresponding Riemannian metric  $d$ , a natural generalization of the  $l^p$ -costs are the costs  $c(\mathbf{x}, \mathbf{y}) = d(\mathbf{x}, \mathbf{y})^p$  and one can not expect in general that this cost is smooth even when  $p \neq 1$ . For example if  $M$  is the sphere  $\mathbb{S}^d$  or the flat torus  $\mathbb{T}^d$  endowed with the Riemannian structure induced by  $\mathbb{R}^{d+1}$ , then  $(\mathbf{x}, \mathbf{y}) \mapsto d(\mathbf{x}, \mathbf{y})^p$  is never differentiable in  $\mathbf{x}$  at  $(\mathbf{x}, \mathbf{y})$  whenever  $\mathbf{x}$  and  $\mathbf{y}$  are antipodal. Nevertheless we shall see in section §5 below that those particular costs also satisfy (NLT) as particular cases of costs associated to a Tonelli Lagrangian.

In view of the previous discussion, one may wonder whether (LT) implies (NLT), or at least if the fact that  $c$  satisfies (LT) on  $U \times N$  implies that it satisfies (NLT) at least on some part of the non-differentiability points for  $c$ . We hereafter provide a counter-example.

**Example 2.9.** The function  $c(x, y) = x^2 y + |x|$  satisfies (LT) on  $\mathbb{R}^2$  but does not satisfy (NLT) : one has  $\nabla_x^- c(0, \cdot) = [-1, 1]$  on  $\mathbb{R}$ .

### 3. A property of transport plans

We now discuss the regularity assumption we shall make upon the initial measure  $\mu$ . As noted in the introduction, a common hypothesis on  $\mu$  is to assume that it does not give mass to sets with  $\mathcal{H}^{d-1}$   $\sigma$ -finite measure. Following Fathi and Figalli [1], we shall make the slightly more general assumption that the measure  $\mu$  does not give mass to countably  $(d-1)$ -Lipschitz sets. We recall the definition below.

**Definition 3.1.** A set  $E \subset \mathbb{R}^d$  is a  $(d-1)$ -Lipschitz graph if there exists a compact subset  $K$  of  $\mathbb{R}^{d-1}$  and a Lipschitz function  $f : K \rightarrow \mathbb{R}$  such that  $E$  is the graph of  $f$  in  $\mathbb{R}^d$  equipped with an adequate orthonormal basis. A set  $E \subset \mathbb{R}^d$  is countably  $(d-1)$ -Lipschitz if it can be expressed as a countable union of  $(d-1)$ -Lipschitz graphs.

It directly follows from the above definition that a countably  $(d-1)$ -Lipschitz set is of  $\mathcal{H}^{d-1}$   $\sigma$ -finite measure : as a consequence if a Borel probability  $\mu$  on  $\mathbb{R}^d$  does not give mass to sets of  $\mathcal{H}^{d-1}$   $\sigma$ -finite measure then it does not give mass to countably  $(d-1)$ -Lipschitz sets. In the case where  $\mu$  is a Borel probability on a second countable  $\mathcal{C}^1$  manifold  $M$ , we say that it does not give mass to countably  $(d-1)$ -Lipschitz sets if this holds in local charts,

that is if  $\phi_\# \mu$  has this property for any local chart  $(V, \phi)$ . Here again this notion does not depend on the chosen local chart : if  $(V, \phi)$  and  $(W, \psi)$  are two local charts then  $\psi \circ \phi^{-1}$  is a  $\mathcal{C}^1$  diffeomorphism on  $\phi(V \cap W)$  so that it preserves the countably  $(d-1)$ -Lipschitz property.

As the proof of our main result in the manifold setting relies on a localization argument, we shall assume that  $M = \mathbb{R}^d$  in the rest of this section. Our aim here is to prove Lemma 3.4 below, which states that when the initial measure  $\mu$  does not give mass to countably  $(d-1)$ -Lipschitz sets then any transport plan  $\Pi(\mu, \nu)$  is concentrated on a  $\sigma$ -compact subset of  $\mathbb{R}^d \times N$  whose elements are somehow interior points of its support (see Remark 3.5 below). We first recall the following definition from Champion, De Pascale and Juutinen [26].

**Definition 3.2.** Let  $\Gamma$  be a  $\sigma$ -compact subset of  $\mathbb{R}^d \times N$ , for  $y \in N$  and  $r > 0$  we define

$$\Gamma^{-1}(\overline{B(y, r)}) := \pi^x(\Gamma \cap (\Omega \times \overline{B(y, r)})).$$

where  $\pi^x : (x, y) \mapsto x$  is the projection on  $\mathbb{R}^d$ .

In terms of mass transportation, if  $\Gamma$  is a  $\sigma$ -compact subset of the support of a transport plan  $\gamma \in \Pi(\mu, \nu)$ , and if  $\gamma$  is concentrated on  $\Gamma$ , then  $\Gamma^{-1}(\overline{B(y, r)})$  may be interpreted as a pre-image of  $\overline{B(y, r)}$  by  $\gamma$  (restricted to  $\Gamma$ ). Note that the set  $\Gamma^{-1}(\overline{B(y, r)})$  is also  $\sigma$ -compact, and thus is a Borel set.

For  $x \in \mathbb{R}^d$ ,  $\xi \in \mathbb{S}^{d-1}$  and  $\delta \in ]0, 1[$  we define the cone

$$C(x, \xi, \delta) := \left\{ y \neq x : \frac{y - x}{\|y - x\|} \cdot \xi > 1 - \delta \right\}, \quad (4)$$

and for a subset  $A$  of  $\mathbb{R}^d$  and  $\varepsilon > 0$  we set

$$C(A; \xi, \delta, \varepsilon) := \{x \in A : C(x, \xi, \delta) \cap A \cap B(x, \varepsilon) = \emptyset\}. \quad (5)$$

The set  $C(x, \xi, \delta)$  is a pointed cone with apex  $x$ , with direction  $\xi$  and angle  $\arccos(1 - \delta)$  around  $\xi$ . Note that when  $A$  is closed, the set  $C(A; \xi, \alpha, \varepsilon)$  is also closed.

The following Lemma of Geometric Measure Theory was shown to us by Tapio Rajala.

**Lemma 3.3.** Let  $A$  be a compact subset of  $\mathbb{R}^d$ ,  $\xi \in \mathbb{S}^{d-1}$  and  $\delta \in ]0, 1[$ , then  $C(A; \xi, \alpha, \varepsilon)$  is a countably  $(d-1)$ -Lipschitz set.

PROOF. Without loss of generality we assume that  $\xi = e_d$  is the  $d$ -th vector of the canonical basis of  $\mathbb{R}^d$ , and we shall denote by  $x_{-d}$  the projection of a vector  $x \in \mathbb{R}^d$  onto  $\mathbb{R}^{d-1} \times \{0\}$ . We fix  $\delta \in ]0, 1[$  and  $\varepsilon > 0$ , and for  $k \in \mathbb{Z}$  we set

$$B_k := C(A; \xi, \delta, \varepsilon) \cap (\mathbb{R}^{d-1} \times [k(1 - \delta)\varepsilon, (k + 1)(1 - \delta)\varepsilon]).$$

Then  $B_k$  is a compact subset of  $\mathbb{R}^d$ . We also remark that if  $x$  and  $y$  both belong to  $B_k$  then  $x \notin C(y, \xi, \alpha)$  and  $y \notin C(x, \xi, \alpha)$ . Using the definition (4) we thus obtain that if  $x$  and  $y$  both belong to  $B_k$  then

$$|y_d - x_d| \leq (1 - \delta)\|y - x\| \leq (1 - \delta)(|y_d - x_d| + \|y_{-d} - x_{-d}\|)$$

so that

$$|y_d - x_d| \leq \frac{1 - \delta}{\delta} \|y_{-d} - x_{-d}\|.$$

We infer from the last inequality that the compact set  $B_k$  is the graph of the Lipschitz function  $x_{-d} \rightarrow x_d$  defined on the projection of  $B_k$  onto  $\mathbb{R}^{d-1} \times \{0\}$ , identified with  $\mathbb{R}^{d-1}$ . Since  $C(A; \xi, \alpha, \varepsilon) \subset \bigcup_{k \in \mathbb{Z}} B_k$  the proof is complete.

The following result is a finer version of Lemma 5.2 and Proposition 5.4 of Champion, De Pascale and Juutinen [26], in which the assumption that  $\mu$  is absolutely continuous with respect to the Lebesgue measure  $\mathcal{L}^d$  is weakened thanks to the above Lemma 3.3.

**Lemma 3.4.** *Assume that the Borel probability measure  $\mu$  on  $\mathbb{R}^d$  gives zero measure to countably  $(d - 1)$ -Lipschitz sets. Let  $\gamma \in \Pi(\mu, \nu)$ , and  $\Gamma$  be a  $\sigma$ -compact set on which  $\gamma$  is concentrated. Then there exists a Borel subset  $R(\Gamma)$  of  $\Gamma \cap \text{support}(\gamma)$  on which  $\gamma$  is concentrated, and such that for any  $(x, y) \in R(\Gamma)$ ,  $r > 0$ ,  $\xi \in \mathbb{S}^{d-1}$ ,  $\delta \in ]0, 1[$  and  $\varepsilon > 0$  one has*

$$x \in \Gamma^{-1}(\overline{B(y, r)}) \setminus C(\Gamma^{-1}(\overline{B(y, r)}); \xi, \delta, \varepsilon). \quad (6)$$

PROOF. Let  $(y_n)_n$  be a dense sequence in  $\mathbb{R}^d$  and  $(\xi_n)_n$  be a dense sequence in  $\mathbb{S}^{d-1}$ , then for any  $(i, j, k) \in \mathbb{N}^5$  we set

$$B_{i,j,k} := C\left(\Gamma^{-1}\left(\overline{B\left(y_i, \frac{1}{j+1}\right)}\right) \cap [-j, j]^d; \xi_k, \frac{1}{j+1}, \frac{1}{j+1}\right).$$

We remark that if  $(x, y) \in \Gamma$  is such that  $x$  belongs to  $C(\Gamma^{-1}(\overline{B(y, r)}); \xi, \delta, \varepsilon)$  for some  $r, \xi, \delta, \varepsilon$  then  $x$  belongs to  $\bigcup_{i,j,k} B_{i,j,k}$ . We infer from Lemma 3.3 that each compact set  $B_{i,j,k}$  is countably  $(d - 1)$ -Lipschitz so that it has  $\mu$ -measure 0. Then the Borel set

$$R(\Gamma) := \Gamma \cap \text{support}(\gamma) \cap \left[ \left( \mathbb{R}^d \setminus \bigcup_{i,j,k} B_{i,j,k} \right) \times N \right]$$

fulfills the desired property.

**Remark 3.5.** *An element  $x$  of  $\pi^x(R(\Gamma))$  of Lemma 3.4 may be considered to be sort of an interior point of  $\pi^x(\Gamma \cap \text{support}(\gamma))$  in the sense that in any direction  $\xi \in \mathbb{S}^{d-1}$  one may find  $x' \neq x$  in  $\pi^x(\Gamma \cap \text{support}(\gamma))$  such that  $\|x - x'\|$  and  $\|\xi - (x - x')/\|x - x'\|\|$  are as small as desired.*

#### 4. $c$ -monotone transport plans are induced by transport maps

In this section we prove the main result of this paper, namely that under mild assumptions on the initial measure  $\mu$  and the cost function  $c$  it holds that any  $c$ -monotone transport plan is induced by transport map. Let us first recall the notion of  $c$ -monotonicity and  $c$ -cyclical monotonicity.



**Definition 4.1.** A subset  $\Gamma$  of  $M \times N$  is  $c$ -monotone whenever for any  $(x_1, y_1)$  and  $(x_2, y_2)$  in  $\Gamma$  it holds

$$c(x_1, y_1) + c(x_2, y_2) \leq c(x_1, y_2) + c(x_2, y_1).$$

The set  $\Gamma$  is  $c$ -cyclically monotone if for any  $N \geq 2$ , any permutation  $\sigma$  of  $\{1, \dots, N\}$  and any  $N$  elements  $(x_1, y_1), \dots, (x_N, y_N)$  of  $\Gamma$  it holds

$$\sum_{i=1}^N c(x_i, y_i) \leq \sum_{i=1}^N c(x_i, y_{\sigma(i)}).$$

When  $c$  is continuous and  $(\mathbf{P})$  has finite value, it is well known that the fact that the support of a transport plan  $\gamma \in \Pi(\mu, \nu)$  is  $c$ -cyclically monotone is a necessary and sufficient condition for its optimality (see e.g. Pratelli [29], Schachermayer and Teichmann [30]). This is of course not the case for  $c$ -monotonicity, as the simple example below shows.

**Example 4.2.** Let  $M = N = \mathbb{R}^2$  and consider the usual squared Euclidian norm cost  $c(x, y) = \|y - x\|^2$ , then the problem  $(\mathbf{P})$  for the probabilities  $\mu = \nu = \frac{1}{\pi} \mathcal{L}^2 \llcorner B(0, 1)$  obviously has a unique solution  $(id \times id)_\# \mu$  induced by the identity, which is then the only admissible transport plan supported by a  $c$ -cyclically monotone set. On the other hand, if  $R_{\pi/2}$  denotes the rotation of center 0 and angle  $\frac{\pi}{2}$  in  $\mathbb{R}^2$  then it is clear that the transport plan  $(id \times R_{\pi/2})_\# \mu$  is  $c$ -monotone and not optimal for  $(\mathbf{P})$ .

We are now in position to state the main result of this paper.

**Theorem 4.3.** Let  $M$  be a second countable  $C^1$  manifold of dimension  $d$  and let  $N$  be a Polish space,  $\mu$  and  $\nu$  be Borel probabilities respectively over  $M$  and  $N$  and let the cost  $c$  be continuous and locally superdifferentiable in  $x$  locally in  $y$  on  $U \times N$ , where  $U$  is an open set on which  $\mu$  is supported. We assume that  $\mu$  gives zero measure to countably  $(d-1)$ -Lipschitz sets and that  $c$  satisfies **(NLT)**.

If a transport plan  $\gamma \in \Pi(\mu, \nu)$  is concentrated on  $\sigma$ -compact and  $c$ -monotone set  $\Gamma$ , then  $\gamma$  is concentrated on a Borel graph.

**Remark 4.4.** Following the line of the proof of Theorem 4.3 below, a similar result holds whenever  $c$  is subdifferentiable instead of superdifferentiable.

**PROOF (PROOF OF THEOREM 4.3).** We first do the proof in the case  $M = \mathbb{R}^d$ .

From Lemma 3.4 we get that  $\gamma$  is concentrated on a Borel subset  $R(\Gamma)$  of  $\Gamma$  which is  $c$ -monotone and such that (6) holds for any  $(x, y) \in R(\Gamma)$ ,  $r > 0$ ,  $\xi \in \mathbb{S}^{d-1}$ ,  $\delta \in ]0, 1[$  and  $\varepsilon > 0$ . We prove by contradiction that  $R(\Gamma) \cap U \times N$  is a graph : assume that  $(x_0, y_0)$  and  $(x_0, z_0)$  both belong to that set with  $y_0 \neq z_0$ . By **(NLT)** we may assume that there exists  $p_0 \in \nabla_x^+ c(x_0, y_0) \setminus \nabla_x^+ c(x_0, z_0)$ , then since the convex set  $\nabla_x^+ c(x_0, z_0)$  is non-empty and closed (see Lemma 2.6), we infer by the Banach separation Theorem that there exists  $\xi \in \mathbb{S}^{d-1}$  and  $\beta > 0$  such that

$$\forall q \in \nabla_x^+ c(x_0, z_0), \quad \xi \cdot (q - p_0) \geq \beta > 0.$$

Now there exists  $\eta : \mathbb{R}^d \rightarrow \mathbb{R}$  such that  $\eta(r) \rightarrow 0$  as  $r \rightarrow 0$  and for all  $(x, z)$  in a neighborhood of  $(x_0, z_0)$  and any  $q_{x,z} \in \nabla_x^+ c(x, z)$  we have

$$c(x_0, z) \leq c(x, z) + q_{x,z} \cdot (x_0 - x) + \eta(x_0 - x) \|x - x_0\|$$

and

$$c(x, y_0) \leq c(x_0, y_0) + p_0 \cdot (x - x_0) + \eta(x - x_0) \|x - x_0\|.$$

We thus obtain

$$c(x_0, z) + c(x, y_0) \leq c(x, z) + c(x_0, y_0) + \|x - x_0\| \left[ \frac{x - x_0}{\|x - x_0\|} \cdot (p_0 - q_{x,z}) + \eta(x - x_0) \right].$$

Since  $(x, z) \mapsto \nabla_x^+ c(x, z)$  has closed graph (see Lemma 2.6) there exist  $\varepsilon > 0$  and  $r > 0$  such that for all  $x \in B(x_0, \varepsilon)$  and  $z \in B(z_0, r)$  one has

$$\forall q_{x,z} \in \nabla_x^+ c(x, z), \exists q \in \nabla_x^+ c(x_0, z_0), \quad \|q_{x,z} - q\| < \frac{\beta}{4}$$

We may also assume that for such a choice of  $\varepsilon$  it holds  $\eta(x - x_0) < \frac{\beta}{4}$ .

Define  $\Delta = \max\{\|q - p_0\| : q \in \nabla_x^+ c(x_0, z_0)\}$  and set  $\delta = [\beta/(8\Delta)]^2$ . Since  $(x_0, z_0)$  is such that (6) holds, we know that there exists  $(x, z) \in \Gamma$  such that  $x \in C(x, \xi, \delta) \cap B(x_0, \varepsilon)$  and  $z \in B(z_0, r)$ . Then  $(x, z)$  satisfy

$$c(x_0, z) + c(x, y_0) \leq c(x, z) + c(x_0, y_0) - \frac{\beta}{4} \|x - x_0\|$$

which contradicts the fact that  $(x, z)$  and  $(x_0, y_0)$  belong to the  $c$ -monotone set  $\Gamma$ .

In the general case where  $M$  is a second countable  $\mathcal{C}^1$  manifold, we consider a  $\mathcal{C}^1$  atlas  $(U_i, \phi_i)_{i \in \mathbb{N}}$  of  $M$  and for any  $i$  we define the cost  $\tilde{c}_i : \phi_i(U_i \cap U) \times N \rightarrow \mathbb{R}$  by  $\tilde{c}_i = c \circ (\phi_i^{-1} \otimes id)$ . For any  $i$  we also associate a compact covering family  $(K_j^i)_j$  of  $U_i \cap U$ , and then for any  $i, j$  we define the initial measure  $\tilde{\mu}_{i,j} = (\phi_i|_{K_j^i})_{\#} \mu$  and the transport plan  $\tilde{\gamma}_{i,j} := ((\phi_i|_{K_j^i}) \times id)_{\#} \gamma$ . Then we are in position to apply the first part of the proof to each triple of data  $(\tilde{c}_i, \tilde{\mu}_{i,j}, \tilde{\gamma}_{i,j})$ , and then obtain that each  $\tilde{\gamma}_{i,j}$  is concentrated on a graph, which yields that  $\gamma$  itself is concentrated on a graph.

Let us now comment the statement of Theorem 4.3 and its proof. In the more recent related results for the existence of an optimal transport map for  $(\mathbf{P})$  in the literature (see Fathi and Figalli [1], Figalli [25, 24], Villani [28]), the basic common assumptions on the data are the following :

- (a) the initial measure  $\mu$  gives zero mass to sets with  $\sigma$ -finite  $\mathcal{H}^{d-1}$  measure;
- (b) the cost  $(x, y) \mapsto c(x, y)$  is lower semicontinuous and locally superdifferentiable in  $x$  locally in  $y$  (generally expressed in terms of local semiconcavity, which is equivalent by Proposition 10.12 in Villani [28]);
- (c) the cost function  $c$  satisfies the Left Twist condition **(LT)**;

(d) the transport plan  $\gamma$  is concentrated on a  $c$ -cyclically monotone set.

In the statement of Theorem 4.3, hypothesis (a) is replaced by the slightly weakened fact that  $\mu$  should give zero measure to countably  $(d-1)$ -Lipschitz sets, which is sufficient for our purpose from Lemma 3.4. Regarding hypothesis (b), in our framework the continuity of  $c$  yields Lemma 2.6, which is necessary to obtain that the superdifferential of  $c$  has closed graph; one may note that continuity of the cost function is in fact a usual feature in this setting, and holds for example for any cost associated with a Tonelli Lagrangian (see Proposition 5.3 below). Then the main differences between the hypotheses of Theorem 4.3 and the common assumptions listed above are that (c) is clearly strengthened while hypothesis (d) is quite weakened, so we shall focus on those two points. As commented after Definition 4.1, under the hypothesis (d) the plan  $\gamma$  is an optimal solution of problem  $(\mathbf{P})$  (or a generalized optimal plan in the case where the value of  $(\mathbf{P})$  is not finite) : as such, its support is included in the  $c$ -subdifferential of a  $c$ -convex function  $\psi$ , which is heuristically a Kantorovich potential for problem  $(\mathbf{P})$  (that is a solution of its dual problem, when the value of  $(\mathbf{P})$  is finite). Then it remains to obtain that this function  $\psi$  is in fact differentiable  $\mu$ -almost everywhere to get that the plan  $\gamma$  is concentrated on the graph of  $D\psi$  : this last point then relies on a deep study of the regularity theory for  $c$ -convex functions. We refer to Chapter 10 of Villani [28] for a presentation of this approach and extended bibliographical notes on this topic. As a consequence, the usual approach to obtain that a plan  $\gamma$  is concentrated on a graph relies on a dual argument and the regularity theory for  $c$ -convex functions. On the contrary, our approach is essentially primal since we assume that the plan  $\gamma$  is concentrated on a set which is just  $c$ -monotone and thus may be not optimal even when the value of problem  $(\mathbf{P})$  is finite (remind Example 4.2). Therefore we can not use the regularity theory for  $c$ -convex functions, and thus the assumption that  $c$  satisfies the Nonsmooth Left Twist condition **(NLT)** compensates the weakening of hypothesis (d) in the statement of Theorem 4.3. As a conclusion, we obtain a result which is more general since it applies to plans that are not necessarily optimal, but also weaker in the sense that we only obtain that such a plan  $\gamma$  is concentrated on the graph of a Borel function (e.g. see Proposition 2.1 in Ambrosio [31]) but we do not get a description of that transport map as a gradient (although one can find such a description as a partial derivative of an Hamiltonian in the special case where  $c(x, y) = \|y - x\|^2$  on  $\mathbb{R}^2$  in Ghoussoub and Moradifam [32]).

## 5. On costs induced by a Lagrangian

In this section we prove that the most common Lagrangian costs do satisfy the Nonsmooth Left Twist condition **(NLT)**. We first recall some basic definitions and known facts about those Lagrangian costs, and refer to Appendix B of Fathi and Figalli [1] and Chapter 10 of Villani [28] for a more detailed presentation.

In all this section  $M$  is assumed to be a connected complete and smooth Riemannian manifold of dimension  $d$ .

**Definition 5.1.** *Given a Lagrangian  $L : TM \rightarrow \mathbb{R}$  (i.e.  $L$  is continuous and bounded by below), to every  $t > 0$  one can associate the cost  $c_t$  on  $M \times M$  defined as*

$$c_t(\mathbf{x}, \mathbf{y}) = \inf \left\{ \int_0^t L(\xi(s), \dot{\xi}(s)) ds : \xi \text{ abs. continuous}, \xi(0) = \mathbf{x}, \xi(t) = \mathbf{y} \right\}.$$

When it exists, a minimizer  $\xi$  of  $c_t(x, y)$  is called an  $L$ -minimal curve.

Some conditions are needed on the Lagrangian  $L$  to ensure finite values and some regularity properties for the costs  $c_t$ . The notions of Tonelli Lagrangian and weak Tonelli Lagrangian are the most common in the Calculus of Variations, the Dynamical System Theory and in Optimal Transportation.

**Definition 5.2.** We say that  $L : TM \rightarrow \mathbb{R}$  is a weak Tonelli Lagrangian on  $M$  if

- (a)  $L$  is  $C^1$  and  $L(\mathbf{x}, \cdot) : T_{\mathbf{x}}M \rightarrow \mathbb{R}$  is strictly convex for any  $\mathbf{x} \in M$ ;
- (b) there exist a complete Riemannian metric  $g$  on  $M$  and a constant  $\alpha > -\infty$  such that

$$\forall (\mathbf{x}, v) \in TM, \quad L(\mathbf{x}, v) \geq \|v\|_{\mathbf{x}} + \alpha$$

where  $\|\cdot\|_{\mathbf{x}}$  is the norm on  $T_{\mathbf{x}}M$  associated to the Riemannian metric  $g$ ;

- (c) for every compact subset  $K \subset M$  the restriction of  $L$  to  $T_K M = \bigcup_{\mathbf{x} \in K} T_{\mathbf{x}}M$  is superlinear in the fibers, that is for every  $a \geq 0$ , there exists a constant  $\alpha(a, K) > -\infty$  such that

$$\forall (\mathbf{x}, v) \in T_K M, \quad L(\mathbf{x}, v) \geq a\|v\|_{\mathbf{x}} + \alpha(a, K).$$

We say that  $L$  is a Tonelli Lagrangian if condition (a) is strengthened in

- (a')  $L$  is  $C^2$  and  $\frac{\partial^2 L}{\partial v^2}(\mathbf{x}, v)$  is positive definite on  $T_{\mathbf{x}}M$  for any  $(\mathbf{x}, v) \in TM$ .

The above notions are sufficient to obtain the following result, see Theorems B.6-7-19 of Fathi and Figalli [1].

**Proposition 5.3.** Let  $L$  be a weak Tonelli Lagrangian and let  $c_t$  be the associated cost for some  $t > 0$ . Then

1.  $c$  is continuous and locally superdifferentiable in  $\mathbf{x}$  locally in  $\mathbf{y}$  on  $M \times M$ ;
2. For any two points  $\mathbf{x}_0$  and  $\mathbf{y}_0$  in  $M$  there exists at least a  $L$ -minimal curve, every such curve  $\gamma_0$  is of class  $C^1$  and moreover

$$-\frac{\partial L}{\partial v}(\mathbf{x}_0, \dot{\gamma}_0(0)) \in \nabla_{\mathbf{x}}^+ c(\mathbf{x}_0, \mathbf{y}_0).$$

We now turn to the main result of this section, namely that under mild assumptions on the weak Tonelli Lagrangian  $L$  it holds that any of the costs  $c_t$  satisfies **(NLT)**.

**Theorem 5.4.** Let  $L : TM \rightarrow \mathbb{R}$  be a weak Tonelli Lagrangian and let  $c_t$  be an associated cost. Assume in addition that  $L$  satisfies the following:

- (UC) If  $\zeta$  and  $\xi$  are two  $L$ -minimal curves with  $\zeta(0) = \xi(0)$  and  $\dot{\zeta}(0) = \dot{\xi}(0)$  then  $\zeta(t) = \xi(t)$ .
- (CP) If  $K \subset M \times M$  is compact then the set of  $L$ -minimal curve  $\{\xi : [0, t] \rightarrow M : (\xi(0), \xi(t)) \in K\}$  is precompact in  $C^1$ .

Then  $c_t$  satisfies **(NLT)**

PROOF. Assume that  $\nabla_{\mathbf{x}}^+ c(\mathbf{x}_0, \mathbf{y}_0) = \nabla_{\mathbf{x}}^+ c(\mathbf{x}_0, \mathbf{z}_0)$ , we aim to show that  $\mathbf{y}_0 = \mathbf{z}_0$ .

First we prove that

$$\nabla_{\mathbf{x}}^+ c(\mathbf{x}_0, \mathbf{y}_0) = \text{conv} \left( \left\{ -\frac{\partial L}{\partial v}(\mathbf{x}_0, \dot{\xi}_0(0)) : \xi \text{ is } L\text{-minimal from } \mathbf{x}_0 \text{ to } \mathbf{y}_0 \right\} \right). \quad (7)$$

By (2) of Proposition 5.3 we already know that

$$\text{conv} \left( \left\{ -\frac{\partial L}{\partial v}(\mathbf{x}_0, \dot{\xi}_0(0)) : \xi \text{ is } L\text{-minimal from } \mathbf{x}_0 \text{ to } \mathbf{y}_0 \right\} \right) \subset \nabla_{\mathbf{x}}^+ c(\mathbf{x}_0, \mathbf{y}_0).$$

If equality does not hold in the previous inclusion then there exist  $p_0 \in \nabla_{\mathbf{x}}^+ c(\mathbf{x}_0, \mathbf{y}_0)$  and  $w \in T_{\mathbf{x}_0} M$  such that

$$p_0 \cdot w < \min \left\{ -\frac{\partial L}{\partial v}(\mathbf{x}_0, \dot{\xi}_0(0)) \cdot w : \xi \text{ is } L\text{-minimal from } \mathbf{x}_0 \text{ to } \mathbf{y}_0 \right\}. \quad (8)$$

Now let  $X_w$  be the geodesic such that  $X_w(0) = \mathbf{x}_0$  and  $\dot{X}_w(0) = w$ , and for  $\varepsilon$  small enough set  $\mathbf{x}_\varepsilon = X_w(\varepsilon) = \exp_{\mathbf{x}_0}(\varepsilon w)$  and  $w_\varepsilon = \dot{X}_w(\varepsilon)$ , where  $\exp_{\mathbf{x}_0}$  denotes the exponential mapping on  $T_{\mathbf{x}_0} M$ . Note that  $X_w^\varepsilon : t \mapsto X_w(\varepsilon - t)$  is also a geodesic and then  $\exp_{\mathbf{x}_\varepsilon}(-\varepsilon w_\varepsilon) = \mathbf{x}_0$ , and also that  $w_\varepsilon \rightarrow w$  as  $\varepsilon \rightarrow 0$  since the geodesic  $X_w$  is smooth. If  $\xi_\varepsilon$  denotes a  $L$ -minimal curve from  $\mathbf{x}_\varepsilon$  to  $\mathbf{y}_0$ , then it follows from (2) of Proposition 5.3 that

$$\frac{\partial L}{\partial v}(\mathbf{x}_\varepsilon, \dot{\xi}_\varepsilon(0)) \cdot (-\varepsilon w_\varepsilon) + o(\varepsilon) \leq c(\mathbf{x}_\varepsilon, \mathbf{y}_0) - c(\mathbf{x}_0, \mathbf{y}_0) \leq p_0 \cdot \varepsilon w + o(\varepsilon).$$

Dividing by  $\varepsilon$  and then passing to the limit, we obtain from the continuity of the Legendre transform and assumption **(CP)** that

$$-\frac{\partial L}{\partial v}(\mathbf{x}_0, \dot{\xi}_0(0)) \cdot w \leq p_0 \cdot w$$

for some  $L$ -minimal curve  $\xi_0$  from  $\mathbf{x}_0$  to  $\mathbf{y}_0$  : this contradicts (8).

We are now in position to conclude from the following fact : when  $A \subset \mathbb{R}^d$  is compact it contains the set of extremal points of  $\text{conv}(A)$  which is non empty by Krein-Milman theorem. By **(CP)** the sets

$$\left\{ -\frac{\partial L}{\partial v}(\mathbf{x}_0, \dot{\xi}_0(0)) : \xi \text{ is } L\text{-minimal from } \mathbf{x}_0 \text{ to } \mathbf{y}_0 \right\}$$

and

$$\left\{ -\frac{\partial L}{\partial v}(\mathbf{x}_0, \dot{\xi}_0(0)) : \xi \text{ is } L\text{-minimal from } \mathbf{x}_0 \text{ to } \mathbf{z}_0 \right\}$$

are compact and by (7) they have same convex hull, so they have nonempty intersection. By the injectivity of  $\frac{\partial L}{\partial v}(x_0, \cdot)$  and by **(UC)** we conclude that  $y_0 = z_0$ .

The following two examples show that the most common Lagrangian costs, and in particular the powers (strictly larger than 1) of the distance on a Riemannian manifold, fall in the framework of our study.

**Example 5.5.** *If  $L$  is a Tonelli Lagrangian then it satisfies the assumptions of Theorem 5.4. Indeed, property (UC) follows by Lemma B.22 and Proposition B.23 of Fathi and Figalli [1], while (CP) is a consequence of the classical regularity theory in the Calculus of Variations. In particular if  $g$  is the Riemannian metric over  $M$  then the square distance cost  $c_t(\mathbf{x}, \mathbf{y}) = t d_g(\mathbf{x}, \mathbf{y})^2$  is associated to the Lagrangian*

$$L(x, v) = g(v, v)$$

*which is a Tonelli Lagrangian.*

**Example 5.6.** *As a corollary of the above example, one also obtains that the costs associated to the Lagrangian*

$$L(x, v) = g(v, v)^{\frac{p}{2}}$$

*for  $p > 1$  also enter the framework of our study : it follows from Proposition B.24 in Fathi and Figalli [1]) that the associated cost for such a Lagrangian is of the form  $c_t(\mathbf{x}, \mathbf{y}) = t^{p-1} d_g(\mathbf{x}, \mathbf{y})^p$ , and as such it has the same  $L$ -minimal curves as the square distance cost, so that the hypotheses of Theorem 5.4 hold. Note that in the case  $p \geq 2$  such a Lagrangian is in fact a Tonelli Lagrangian.*

## Acknowledgments

The first author wishes to thank the Equipe de Combinatoire et Optimisation (Institut de Mathématiques de Jussieu, CNRS UMR 7586, Université Pierre et Marie Curie, France) where part of this work was done.

The research of the second author is part of the project 2008K7Z249 “*Trasporto ottimo di massa, disuguaglianze geometriche e funzionali e applicazioni*”, financed by the Italian Ministry of Research and is partially financed by the “*Fondi di ricerca di ateneo*” of the University of Pisa.

Visits of both authors have been partially supported by the Universities of Pisa and Toulon.

- [1] A. Fathi, A. Figalli, Optimal transportation on non-compact manifolds, Israel J. Math. 175 (2010) 1–59. doi:10.1007/s11856-010-0001\_5.  
URL [http://dx.doi.org/10.1007/s11856-010-0001\\_5](http://dx.doi.org/10.1007/s11856-010-0001_5)
- [2] Y. Brenier, Décomposition polaire et réarrangement monotone des champs de vecteurs, C. R. Acad. Sci. Paris Sér. I Math. 305 (19) (1987) 805–808.
- [3] L. Rüschendorf, S. T. Rachev, A characterization of random variables with minimum  $L^2$ -distance, J. Multivariate Anal. 32 (1) (1990) 48–54. doi:10.1016/0047-259X(90)90070-X.  
URL [http://dx.doi.org/10.1016/0047-259X\(90\)90070-X](http://dx.doi.org/10.1016/0047-259X(90)90070-X)

- [4] L. A. Caffarelli, Allocation maps with general cost functions, in: Partial differential equations and applications, Vol. 177 of Lecture Notes in Pure and Appl. Math., Dekker, New York, 1996, pp. 29–35.
- [5] W. Gangbo, R. J. McCann, Optimal maps in Monge’s mass transport problem, C. R. Acad. Sci. Paris Sér. I Math. 321 (12) (1995) 1653–1658.
- [6] W. Gangbo, R. J. McCann, The geometry of optimal transportation, Acta Math. 177 (2) (1996) 113–161. doi:10.1007/BF02392620.  
URL <http://dx.doi.org/10.1007/BF02392620>
- [7] L. Rüschendorf, On  $c$ -optimal random variables, Statist. Probab. Lett. 27 (3) (1996) 267–270. doi:10.1016/0167-7152(95)00078-X.  
URL [http://dx.doi.org/10.1016/0167-7152\(95\)00078-X](http://dx.doi.org/10.1016/0167-7152(95)00078-X)
- [8] L. Kantorovith, On the translocation of masses, C.R. (Dokl.) Acad. Sci. URSS 37 (2) (1942) 199–201.
- [9] L. Kantorovith, On a problem of monge (in russian), Uspekhi Mat. Nauk. 3 (2) (1948) 225–226.
- [10] G. Monge, Mémoire sur la théorie des déblais et des remblais, Histoire de l’Académie des Sciences de Paris.
- [11] L. C. Evans, W. Gangbo, Differential equations methods for the Monge-Kantorovich mass transfer problem, Mem. Amer. Math. Soc. 137 (653) (1999) viii+66.
- [12] L. Ambrosio, A. Pratelli, Existence and stability results in the  $L^1$  theory of optimal transportation, in: Optimal transportation and applications (Martina Franca, 2001), Vol. 1813 of Lecture Notes in Math., Springer, Berlin, 2003, pp. 123–160. doi:10.1007/978-3-540-44857-0\_5.  
URL [http://dx.doi.org/10.1007/978-3-540-44857-0\\_5](http://dx.doi.org/10.1007/978-3-540-44857-0_5)
- [13] L. A. Caffarelli, M. Feldman, R. J. McCann, Constructing optimal maps for Monge’s transport problem as a limit of strictly convex costs, J. Amer. Math. Soc. 15 (1) (2002) 1–26 (electronic). doi:10.1090/S0894-0347-01-00376-9.  
URL <http://dx.doi.org/10.1090/S0894-0347-01-00376-9>
- [14] L. Caravenna, A proof of Sudakov theorem with strictly convex norms, Math. Z. 268 (1-2) (2011) 371–407. doi:10.1007/s00209-010-0677-6.  
URL <http://dx.doi.org/10.1007/s00209-010-0677-6>
- [15] T. Champion, L. De Pascale, The Monge problem in  $\mathbb{R}^d$ : variations on a theme, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) 390 (Teoriya Predstavlenii, Dinamicheskie Sistemy, Kombinatornye Metody. XX) (2011) 182–200, 309.
- [16] T. Champion, L. De Pascale, The Monge problem for strictly convex norms in  $\mathbb{R}^d$ , J. Eur. Math. Soc. (JEMS) 12 (6) (2010) 1355–1369. doi:10.4171/JEMS/234.  
URL <http://dx.doi.org/10.4171/JEMS/234>

- [17] N. S. Trudinger, X.-J. Wang, On the Monge mass transfer problem, *Calc. Var. Partial Differential Equations* 13 (1) (2001) 19–31. doi:10.1007/PL00009922.  
URL <http://dx.doi.org/10.1007/PL00009922>
- [18] C. Villani, Topics in optimal transportation, Vol. 58 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2003. doi:10.1007/b12016.  
URL <http://dx.doi.org/10.1007/b12016>
- [19] G. Carlier, Duality and existence for a class of mass transportation problems and economic applications, in: *Advances in mathematical economics*. Vol. 5, Vol. 5 of *Adv. Math. Econ.*, Springer, Tokyo, 2003, pp. 1–21. doi:10.1007/978-4-431-53979-7-1.  
URL <http://dx.doi.org/10.1007/978-4-431-53979-7-1>
- [20] V. Levin, Abstract cyclical monotonicity and Monge solutions for the general Monge-Kantorovich problem, *Set-Valued Anal.* 7 (1) (1999) 7–32. doi:10.1023/A:1008753021652.  
URL <http://dx.doi.org/10.1023/A:1008753021652>
- [21] D. Cordero-Erausquin, Sur le transport de mesures périodiques, *C. R. Acad. Sci. Paris Sér. I Math.* 329 (3) (1999) 199–202. doi:10.1016/S0764-4442(00)88593-6.  
URL [http://dx.doi.org/10.1016/S0764-4442\(00\)88593-6](http://dx.doi.org/10.1016/S0764-4442(00)88593-6)
- [22] R. J. McCann, Polar factorization of maps on Riemannian manifolds, *Geom. Funct. Anal.* 11 (3) (2001) 589–608. doi:10.1007/PL00001679.  
URL <http://dx.doi.org/10.1007/PL00001679>
- [23] P. Bernard, B. Buffoni, Optimal mass transportation and Mather theory, *J. Eur. Math. Soc. (JEMS)* 9 (1) (2007) 85–121. doi:10.4171/JEMS/74.  
URL <http://dx.doi.org/10.4171/JEMS/74>
- [24] A. Figalli, The Monge problem on non-compact manifolds, *Rend. Semin. Mat. Univ. Padova* 117 (2007) 147–166.
- [25] A. Figalli, Existence, uniqueness, and regularity of optimal transport maps, *SIAM J. Math. Anal.* 39 (1) (2007) 126–137. doi:10.1137/060665555.  
URL <http://dx.doi.org/10.1137/060665555>
- [26] T. Champion, L. De Pascale, P. Juutinen, The  $\infty$ -Wasserstein distance: local solutions and existence of optimal transport maps, *SIAM J. Math. Anal.* 40 (1) (2008) 1–20. doi:10.1137/07069938X.  
URL <http://dx.doi.org/10.1137/07069938X>
- [27] T. Champion, L. De Pascale, The Monge problem in  $\mathbb{R}^d$ , *Duke Math. J.* 157 (3) (2011) 551–572. doi:10.1215/00127094-1272939.  
URL <http://dx.doi.org/10.1215/00127094-1272939>
- [28] C. Villani, Optimal transport, Vol. 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, Springer-Verlag, Berlin, 2009, old and new. doi:10.1007/978-3-540-71050-9.  
URL <http://dx.doi.org/10.1007/978-3-540-71050-9>



- [29] A. Pratelli, On the sufficiency of  $c$ -cyclical monotonicity for optimality of transport plans, *Math. Z.* 258 (3) (2008) 677–690. doi:10.1007/s00209-007-0191-7.  
URL <http://dx.doi.org/10.1007/s00209-007-0191-7>
- [30] W. Schachermayer, J. Teichmann, Characterization of optimal transport plans for the Monge-Kantorovich problem, *Proc. Amer. Math. Soc.* 137 (2) (2009) 519–529. doi:10.1090/S0002-9939-08-09419-7.  
URL <http://dx.doi.org/10.1090/S0002-9939-08-09419-7>
- [31] L. Ambrosio, Lecture notes on optimal transport problems, in: *Mathematical aspects of evolving interfaces* (Funchal, 2000), Vol. 1812 of *Lecture Notes in Math.*, Springer, Berlin, 2003, pp. 1–52. doi:10.1007/978-3-540-39189-0-1.  
URL <http://dx.doi.org/10.1007/978-3-540-39189-0-1>
- [32] N. Ghoussoub, A. Moradifard, A note on simultaneous preconditioning and symmetrization of non-symmetric linear systems, *Numer. Linear Algebra Appl.* 18 (3) (2011) 343–349. doi:10.1002/nla.730.  
URL <http://dx.doi.org/10.1002/nla.730>